Defense Technical Information Center

Compilation Part Notice

ADP021363

TITLE: MKIDS - Management of Knowledge Intensive Dynamic Systems

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: International Conference on Integration of Knowledge Intensive Multi-Agent Systems. KIMAS '03: Modeling, Exploration, and Engineering Held in Cambridge, MA on 30 September-October 4, 2003

To order the complete compilation report, use: ADA441198

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP021346 thru ADP021468

MKIDS - Management of Knowledge Intensive Dynamic Systems

Sidney Berkowitz, Ph.D.
Department of Defense
Fort Meade, Maryland 20755
s.berkowitz@ieee.org

Abstract - *MKIDS is a research program sponsored by the National Science Foundation (NSF) for the Department of Defense (DOD). The purpose of the program is to create management tools and methodologies that permit the scheduling and control of production processes which are knowledge intensive and react to dynamic changes in the external and operating environments with agility. NSF recognizes that this research will be of interest to wide sectors of emerging commerce as well as DOD production assets. This paper will describe the MKIDS paradigm. The papers presented in this KIMAS '03 session will deal with the scheduling, social metrics, and organizational behavior of the paradigm.*

1. INTRODUCTION

Industry, e-commerce among others, has embraced the idea of tightly integrating statistical data mining and information retrieval techniques within its decision support -systems. These techniques work most favorably for business environments that are stationary, for processes that are rigid, and for information needs that are fairly narrow in scope (e.g. How many clicks are recorded...). Currently, managers of our leading commercial enterprises are well disposed to enlist performance history as a decision indicator for the future. As we move into knowledge intensive ventures that depend on quick market reaction times for survival, the increasingly dynamic requirements of doing business will create a corresponding decrease in the dependence on long-term experience and deep capitalization. Unlike the manufacturing sector, which bears the built-in inertia peculiar to physical plants, a knowledge based entity is especially well suited for the rapid re-configuration required to meet a perceived market need. In a physical plant, re-configuration is difficult because of the fixed nature of machines to which humans are merely adjunct assistants. In a knowledge based industry, on the other hand, people are the principal production elements, to whom machines are merely

assistants with formidable memory but scant intelligence. In such an industry, the firm can adapt to a changing environment with all the flexibility of creative minds. On the other hand, the experience of the select domain expert is likely not to suffice in a world where the extent, growth, and change of knowledge is unlikely to be within the grasp of any one individual, who, in any event, is not likely to remain with any one company long enough to acquire much more than a modest measure of expertise. Rather, we see, or hope for, teams of individuals supported by computerized knowledge bases, collaborating on well-informed decision-making. We are poised to enter a new phase of management practice in which structured, global collections of knowledge will be brought to bear in assisting managers with complex, rapid decision making. This knowledge management capability will have the effect of diffusing expertise throughout the corporate structure, allowing for flexible process configuration and stimulating an intelligent management response to external change. For example, we will be able to build on demand, gracefully adjust to the failure of local component processes, and take into account the effects of foreign policies and politics on our local enterprise. We expect such benefits all without maintaining cadres of domain experts whose lives are devoted to pursuing our queries or maintaining watch for events of interest.

2. THE MKIDS PARADIGM

Definitions

We address the problem of managing those business processes whose *dynamics* are sensitive to a *complex of knowledge* both within the process and throughout the surrounding environment.

By *dynamics*, we mean the rapid re-configuration or re-scheduling of resources, human or machine, in response to events that occur either within or outside the process. Internal events might include data or information that describe the performance of a process, measurements reflecting the deterioration of a component operation, historical workflow, or the status of inventory. External events might include a political crisis, market shift, emergence of new sources of material. In general, the events may require a reconfiguration, reallocation of assets,
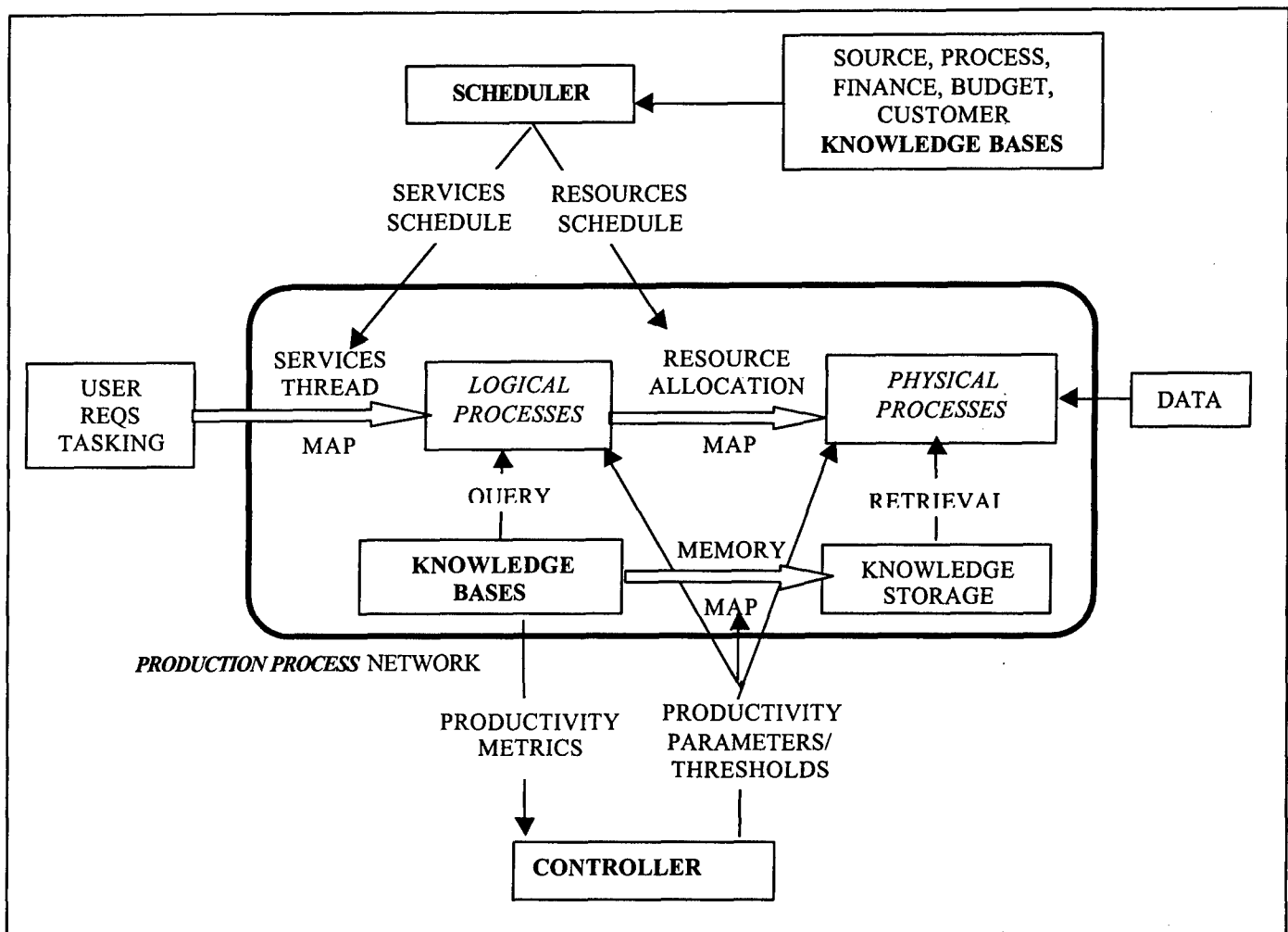
Figure 1 - Management paradigm schematic of knowledge intensive production process

or reprioritization of tasks. Thus, Merrill Lynch, Verizon, and FedEx are examples of companies that exhibit resource allocation and scheduling dynamics. However, the operation of Verizon does not critically depend on a conceptual knowledge base, as we define it below, and, therefore, does not pass the bar of our problem requirement. In general, the complexity of a network topology, together with its switching and error-detection functionality, is not the problem area we are addressing. We are looking at the dynamics of a knowledge intensive industry, as we now discuss.

By *complex of knowledge*, we mean the set of domain specific knowledge bases that are peculiar to a given business. Thus, the World Bank, the Associated Press, Texas Instruments, and the Patent and Trademarks Office (PTO) are examples of organizations that are knowledge-based - although the current knowledge reservoir may be mostly human. The PTO, however, has a moderately rigid production workflow; its process dynamics are not within the scope of the proposed area of study. The domain knowledge bases of these organizations include concept graphs in areas ranging from banking policy to geo-political

perspective to technical design knowledge. These knowledge bases might make reference to numeric or text data bases, but, even so, the knowledge bases claim a level of abstraction that structured collections of data/information do not possess. Thus, web-based companies like Google.com or Amazon.com depend on huge data bases of numbers or text, perhaps on profile data derived from collaborative filters - as in market-basket analysis. However, aside from some crude thesaurus - a simple, universal ontology -, these companies do not currently require abstract knowledge to run their operations. Most of all, even if current business reasoning is based on data mining or information retrieval techniques, business decisions and inferences are not yet drawn or assisted by computer from an intelligent view of some slice of the real world.

Description of knowledge intensive dynamic systems

The schematic in Figure 1 is a naïve representation of the scheduling and control mechanisms for a production process. In the schematic, there are two major management functions depicted: the **scheduler** and the **controller**. The

101

scheduler maps user requirements for processing data into task threads strung through a logical process network. For example, a thread for a news distribution organization might consist of newsgathering, triage, rewriting, editing, proofreading, layout, publishing, and dissemination functions. The **scheduler** then maps the logical threads into sequences of tasks on physical resources, including, for example: communications links, software filtering and editing services, computers, and the most important resource of all: people. The **scheduler** decisions are supported by an array of management knowledge sources, distinct from the production knowledge sources that support the logical process network. The management knowledge sources include knowledge of data sources, the state of the production process, finance and budget, and customer profiles. During execution of the production process, the **controller** measures both local and global productivity and attempts to tune the process productivity by:

- adjusting parameters and thresholds of the logical process models
- re-configuring the physical process network
- re-mapping process knowledge in a distributed knowledge storage network.

The **controller** also affects productivity through managing functions such as training, customer service, and human relations, which only indirectly relate to the production process. In principle, the effects of these functions, too, could be a source of productivity models within the production process.

For large, complex processes, a centralized scheduling strategy for tasks in a large production process is likely to congest due to interdependencies and priority conflicts in the task queues. At large scale, the physical resource allocation time increases non-linearly with queue length. A large cost is due to the communications burden that is incurred by the need to know and modify the queue status. The introduction of new tasks causes a ripple effect in re-prioritizing the task queues. Introducing new physical resources or accounting for outages incurs an ongoing life cycle cost in scheduling process maintenance. The flip side of this coin is a completely distributed mobile software agent strategy which is apparently scalable. We buy this flexibility at the cost of thrashing or bottlenecks due to greedy, myopic planning in which the local planner does not have contextual state information. That is, the local planner lacks competence to plan effectively because it does not have a corporate viewpoint. As a practical matter, there is still a communications burden due to the responsibility of the production process to track the status of tasks and report on that status to the customer. We need some compromise design scheme to defeat the scale problem, perhaps a hierarchical scheduler in which planning authority resides at various levels of the hierarchy with a corresponding

broadened view of processing needs. Even with a scheduling solution in hand, we need to understand the limitations of the solution for a given capitalization.

The control strategy and the scheduling strategy are complementary in the sense that control decisions effectively modify data queues or flows by re-parameterizing or re-thresholding process models. These changes then reflect a modified task ranking or selection in the scheduling process. The reason for this tasking modification is that productivity is not simply a measure of transport flow, but is also a measure of the quality and perhaps the timeliness of the intermediate products being processed. That is, the production process is not simply a communications network which passes material unchanged from one end of the network to the other. Rather, it is a business process network at each node of which a sometimes sophisticated, knowledge-based transformation of material takes place, as we discuss in the next management issue. Since the transformations may be interdependent, it becomes increasingly difficult, as the scale of the processing network increases, to decide how to fuse local productivity measures and how to optimize global productivity. Perhaps we need operational simulations to help us search the space of scheduling and control decisions.

In general, **schedulers** and **controllers** may be a distributed complex of humans and machines. Scheduling and control strategies are the ultimate products of MKIDS. They are difficult to define effectively just because the production process is knowledge intensive and dynamic, as we now discuss.

3. MANAGEMENT ISSUES

Issue: Complexity of knowledge processing productivity

At various points in a production process thread, humans or machines access stored knowledge, apply that knowledge to the product being processed, and perhaps evaluate the quality, relevance, and/or next destination of the product. For example, an editor might critique the style of a news report re-write and send it back for another go, or send it forward with a layout assignment to the back page because of the lack of current significance. Yet, capturing the implicit productivity measurement of either the editor or the reporter who submitted the report is a difficult thing to do from both a technical and a cultural viewpoint. How shall we define productivity at each knowledge-processing node? How shall we instrument the workplace to measure productivity? How shall we express the dependencies of a productivity metric on prior services in the thread? A critical aspect of accessing knowledge is that knowledge may not be obviously available. How do we then manage the sharing of knowledge stored in distributed, incompatible databases, or, worse, in peoples' heads, in fleeting

telephone/hallway conversations, formal meetings, or on yellow stickies? Can we induce and grow domain knowledge from any of these sources? Can the process automatically or interactively understand a functional requirement well enough to contribute collaborative knowledge on its own? Not least, how might we motivate the workforce to participate overtly in knowledge sharing and in the relevance evaluation procedure, or shall we measure functional productivity implicitly? People may not be ready for Big Brother, but they may also see their active participation in quality assessment as intrusive or burdensome. How do we design and implement a control policy to encourage high productivity? Might we offer the opportunity to process more data or to get better assignments?

Issue: Hidden workflow

In modeling the production process for scheduling and control purposes, the workflow rules that map user requirements to product are often informal, implicit to the process, not stationary, and do not necessarily reflect the organizational hierarchy. Scheduling becomes subject to the priority of the moment. Measuring and controlling productivity in such a situation is a daunting challenge. After all, if the workflow rules are not explicit, what process shall we measure? Moreover, if rules are not conditioned on snapshot events but rather on sequences of prior events, how shall we limit our search of the sequence space for appropriate triggers, a potentially combinatorial experience? In general, maintenance of temporal sequences in a knowledge base is a poorly understood problem.

An even more disturbing aspect of not understanding how business is actually conducted is that knowledge that is shared during hidden transactions is highly volatile. Even if the process requires that operational logs be kept, the logs are often incomplete and inconsistent. Hidden knowledge sharing has some unfortunate consequences:

- domain knowledge is not universally available to those who might need it
- re-creating such knowledge is a gross process inefficiency
- people require long lead times to learn new work roles
- unvalidated knowledge tends to be corrupt

Is it possible to create a background process that learns or infers workflow rules and missing knowledge based on observed behavior reflected, say, in operational logs? Can we successfully require or encourage people to record and share knowledge? Acquisition of workflow rules depends on some quasi-stationary regularity in the process. If the process continually changes, can we devise a learning

process that will keep pace? On the other hand, will we be able to capture rules that are triggered by rare events?

Issue: Business process dynamics

Changes in markets or sources of material, build-to-order strategies, new technology, outages, tax structure, legal constraints and the like may provoke a need to re-configure or at least re-schedule the production process frequently. When this happens, having a domain knowledge expert can act as a stabilizing force to ease the transition from the old to the new process. Such changes require an adaptive control strategy that tracks productivity measures, detects changes, and appropriately re-builds process models.

4. CONCLUSION: MKIDS RESEARCH AREAS

Having described the MKIDS paradigm and the management issues associated with that paradigm, we now list the following basic research areas related to the management of knowledge intensive business dynamics.

1. Adaptive scheduling and control of product dynamics – We need to find methods for solving a multi-objective optimization problem in a dynamically changing environment, with the following characteristics. As the business process transforms source material and knowledge into product, resources must be allocated, perhaps shared, and dynamically re- allocated to meet productivity goals and priorities. Of special importance is the fact that goals and priorities themselves change rapidly in response to rarely seen and therefore hard-to-predict events: market shifts, technical breakthroughs, geo-political conflict. As information makes its way through the process, component functions must triage, sift, and co-ordinate the information. Optimization of the flow might involve more than maximization of quantity; it might mean optimization of quality. Dynamic tuning can take place by measuring productivity at the local component level and/or by monitoring global patterns of network activity. In this regard, we need to understand the "eigenstructure" of the business process, the key control parameters which determine system behavior.

2. Collaborative knowledge representation, acquisition, retrieval, and inference - We are interested in research in the representation and sharing of management knowledge related to knowledge intensive dynamic business processes. We need to find ways to define, acquire, and store corporate knowledge from which we can glean information relevant to a given process. Very often, we cannot predict derivative knowledge and thereby store it in a structured manner. Therefore, we

need to devise methods of automating the gathering and representation of knowledge about areas that include but are not limited to: source material, design strategies, scheduling methodologies, customer requirement profiles, process status, energy sources, and geo-politics. The resultant set of management domain knowledge bases must support inferencing across the domain set. Indeed, we need to investigate the inference engines themselves within the bounds of this task..

3. Learning workflow rules – Can we discover means of automatically acquiring expertise in scheduling and control by monitoring human decisions or by experimentally - say by simulation - probing the process? Perhaps we do not have to view the previous adaptive control task as an optimization problem, but one of seeking feasible solutions through some learning mechanism. For example, we might investigate reinforcement learning techniques that balance greedy against exploratory searches of the solution space. Alternatively, we might instead search the search strategy space by an evolutionary optimization mechanism. The nub of the problem is that we find it difficult, as humans, to formulate *a priori* scheduling and control strategies for complex processes and therefore need a computer-based means of inducing these strategies from observation and experience.

4. Distributed decision-making – What are the consequences of empowering operating levels of the management hierarchy with autonomous decision-making authority? In a complex, flexible business process whose configuration is constantly changing, we may find controlling the enterprise from the top a daunting task. The problem is, in part, a question of how to share information about requirements and resources, and, in part, how to endow the actors, both human and machine, with authority and responsibility. How shall we partition the available information flow and resource tradespace to match the processing need of the work function? Perhaps the function should have the authority to actively bid on resources rather than be assigned them. We need to formulate verifiable methodologies that will empower decision agents throughout the process with limited control over their operational functions and productivity without detracting from the overall yield of the enterprise.

5. Performance metrics and marketing – We think that none of the efforts in adaptive control, learning, or distributed decision-making will succeed without innovative schemes for assessing the productivity of a dynamically re-configuring business process. Monitoring the bottom line is a traditional, myopic way of assessing the adequacy and power of a business operation. Even a favorable snapshot of overall operations is not a guarantee of long-term success in a marketplace that, for example, fluctuates with rapid changes in technology. This critique is sharpened when the statistics that one might use for prediction turn out to be non-stationary. We need to be able to predict, not just performance, but rather customer satisfaction, and to trace that global metric back to local component performance. In general, this is a difficult inverse problem in which we seek to identify the cause(s) of global success or failure by auditing local behavior. We need to create methods for modeling the customer and establishing criteria for the utility of a product. These methods might range from active, intrusive, but nevertheless co-operative, monitoring of customer behavior (e.g. Arbitron, Nielsen polling) to unobtrusive lead-ins (e.g. read Insight for 6 months, then subscribe) to passive advertising (e.g. Citibank is here to help you meet your financial obligations). We even need to create novel means for motivating the acceptance and implementation of metrics in a knowledge production environment, since management by numbers can easily founder on the recalcitrant adherence to a work culture that has been hitherto free of detailed monitoring. Behavior modification and perception management are needed in order to manage by numbers, and these are at best uncertain science. Finally, if we are ever to face reality, metrics have a political or image aspect that must be considered, since exposing customers or investors to misleading measures of performance, however insightful they may be to management, could prove detrimental to the company. Certainly, each business is different but are there useful principles that apply to various types of commerce?

The papers presented in this KIMAS '03 session represent investigations into some of the research issues described above by a selection of MKIDS sponsored researchers.